

Copyright (c) 1993 - 2003 Compugen Ltd.	GenCore version 5.1.3					
Om protein - protein search, using sw model						
Run on:	February 11, 2003, 19:44:49 ; Search time 16.0286 Seconds (without alignments) 2650.980 Million cell updates/sec					
Title:	US-09-497-967-6					
Perfect score:	2342					
Sequence:	1 MKYILLILLLISLNFELRA.....STTPAKFLSISLLFISFYLL 442					
Scoring table:	BLOSUM62					
Gapext:	0.5					
Searched:	283224 seqs, 96134422 residues					
Total number of hits satisfying chosen parameters:	283224					
Minimum DB seq length:	0					
Maximum DB seq length:	2000000000					
Post-processing:	Minimum Match 0% Maximum Match 100%					
Database :	PIR_73;* 1: Piri;* 2: Pir2;* 3: Pir3;* 4: Pir4;*					
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.	283224					
%	Query Score Match Length DB ID Description					
Result No.	Score	Match	Length	DB	ID	Description
1	2083	88.9	395	2	A46031	immobilization sur
2	213.5	9.1	713	2	A35502	major surface-labe
3	212	9.1	124	2	T42017	cysteine rich prot
4	210.5	9.0	557	2	A4434	variant-specific s
5	210.5	9.0	596	2	A45664	variant-specific s
6	206	8.8	1766	2	A42125	trophozoite cystei
7	205	8.8	1459	2	T32271	hypothetical prote
8	203	8.7	677	2	C42125	trophozoite cystei
9	202.5	8.6	3625	2	T10053	laminin alpha 5 ch
10	198	8.5	3084	1	MHMSA	laminin alpha 1 ch
11	195.5	8.3	1372	2	T25933	hypothetical prote
12	189	8.1	5376	2	T42215	zonadhesin - mouse
13	183.5	7.8	1895	2	T15881	hypothetical prote
14	183	7.8	1299	2	T43251	furin (EC 3.4.21.7)
15	182	7.8	667	2	A48579	trophozoite surfac
16	181.5	7.7	3075	2	S14458	laminin alpha-1 ch
17	180.5	7.7	2395	1	S50820	hypothetical prote
18	178	7.6	3106	1	S53868	laminin alpha-2 ch
19	177.5	7.6	1607	1	MMSB2	laminin gamma-1 ch
20	177	7.6	439	2	A36385	surface antigen se
21	173	7.4	1790	1	MMFFB1	laminin beta-1 cha
22	172	7.3	899	2	G0428	subtilisin-like pr
23	172	7.3	915	2	JC6148	subtilisin-like pr
24	170.5	7.3	2195	2	T34264	hypothetical prote
25	169.5	7.2	1169	2	S3181	flocculation prote
26	169	7.2	1548	2	S34583	serine protease
27	166.5	7.1	1557	2	T28811	hypothetical prote
28	166	7.1	1797	2	T21889	hypothetical prote
29	166.5	7.1	1805	2	T21888	hypothetical prote

ALIGNMENTS

RESULT 1	A46031	Immobilization surface I-antigen precursor - Ichthyophthirius multifiliis (fragment)
C; Species:	Ichthyophthirius multifiliis	
C; Date:	29-Sep-1999	#sequence_revision 29-Sep-1999 #text_change 07-Dec-1999
C; Accession:	A46031	
R; Clark, T.G.; McGraw, R.A.; Dickerson, H.W.		
Proc. Natl. Acad. Sci. U.S.A. 89, 6363-6367, 1992		
A; Title:	Developmental expression of surface antigen genes in the parasitic ciliate I	
A; Reference number:	A46031; MUID:92335298; PMID:1631132	
A; Accession:	A46031	
A; Molecular type: mRNA; protein		
A; Residues: 1-395 <CLAs>		
A; Cross-references: GB:M92907; NID:93628568; PID:93628569		
A; Note: the authors translated the codon UUG for residue 330 as Tle		
A; Note: sequence extracted from NCBI backbone (NCBIN:108734, NCBIP:108735); the sequence		
C; Genetics:		
A; Genetics code: SEC5		
C; Keywords: glycoprotein; surface antigen		
F; 2:395/Product: immobilization surface I-antigen #status experimental <PAT>		
F; 156,191,245,281/Binding site: carbohydrate (Asn) (covalent) #status predicted		
Best Local Similarity 99.2%; Pred. No. 4-9e-125; Mismatches 2; Indels 0; Gaps 0;		
Matches 387; Conservative 2;		
Qy 20 AVCPDPGTQAGITDVGADLGTVCNCRPNFYNGAAQEANGNQPFAAANNAARGICCV 79		
Db 1 AVCPDPGTQAGITDVGADLGTVCNCRPNFYNGAAQEANGNQPFAAANNAARGICCV 60		
Qy 80 PCQINRGSVTNAGDIATLATCSPQCPSTALDGTVDFDSAAQCVPCKPNFYNGG 139		
Db 61 PCQINRGSVTNAGDIATLATCSPQCPSTALDGTVDFDSAAQCVPCKPNFYNGG 120		
Qy 140 SPGEAPGVQEAAAGAAAGVAATYSQCPOLNKNDSPATACQAQNIALATOCSNOCPNGT 199		
Db 121 SPQEAPGVQEAAAGAAAGVAATYSQCPOLNKNDSPATACQAQNIALATOCSNOCPNGT 180		
Qy 200 VLDDGTVLYFNTSATLCVKCRPNFYNGGSPQGEAPGVQEAAAGVAATYSQVPC 259		
Db 181 VLDDGTVLYFNTSATLCVKCRPNFYNGGSPQGEAPGVQEAAAGVAATYSQVPC 240		
Qy 260 QINKNDSPATAGAQANLAQTCSTQCPITGAQDGVTLVFSNSSTQCSOCIANYFFNGNE 319		
Db 241 QLNKNDSPATAGAQANLAQTCSTQCPITGAQDGVTLVFSNSSTQCSOCIANYFFNGNE 300		
Qy 320 AGKSQCLKCPVSKTPAHGNTATQATQCLTCTPAQGTYLDDGTSTMFVASATECTKCSA 379		
Db 301 AGKSQCLKCPVSKTPAHGNTATQCLTCTPAQGTYLDDGTSTMFVASATECTKCSA 360		
Qy 380 GFFASKTTGFTAGIDTCTECTKMLTSQATGATA 409		

	A;Title: A Giardia duodenalis gene encoding a protein with multiple repeats of a toxin
	A;Reference number: Z22207
	A;Accession: T42017
	A;Status: preliminary; translated from GB/EMBL/DDBJ
RESULT 2	A;Molecule type: DNA
A35502	A;Cross-references: EMBL:1-1274 <CHE>
major surface-labeled trophozoite antigen precursor - Giardia lamblia	C;Species: Giardia lamblia
C;Date: 09-Nov-1990 #sequence_revision 09-Sep-1997	C;Accession: A35502
C;Sequence: L29079; PID:9951190; PID:AAA74587.1	C;Experimental source: specific host: Homo sapiens
C;Residues: 1-1274	C;Superfamily: unassigned ankyrin repeat proteins
Query Match 9.1%; Score 212; DB 2; Length 1274;	Query Match 9.1%; Score 212; DB 2; Length 1274;
Best Local Similarity 20.9%; Pred. No. 4.4e-06;	Best Local Similarity 20.9%; Pred. No. 4.4e-06;
Matches 97; Conservative 42; Mismatches 169; Gaps 21;	Matches 97; Conservative 42; Mismatches 169; Gaps 21;
A;Title: Isolation and expression of the gene for a major surface protein of Giardia lamblia	A;Title: Isolation and expression of the gene for a major surface protein of Giardia lamblia
A;Reference number: A35502; PMID:9028035; PMID:2352929	A;Reference number: A35502; PMID:9028035; PMID:2352929
A;Accession: A35502	A;Accession: A35502
A;Status: preliminary	A;Status: preliminary
A;Molecule type: DNA	A;Molecule type: DNA
A;Cross-references: GB:M33641; NID:9159131; PID:9159132	A;Cross-references: GB:M33641; NID:9159131; PID:9159132
C;Keywords: surface antigen; transmembrane protein	C;Keywords: surface antigen; transmembrane protein
Query Match 9.1%; Score 213.5; DB 2; Length 713;	Query Match 9.1%; Score 213.5; DB 2; Length 713;
Best Local Similarity 23.0%; Pred. No. 2.2e-06;	Best Local Similarity 23.0%; Pred. No. 2.2e-06;
Matches 127; Conservative 42; Mismatches 187; Indels 31;	Matches 127; Conservative 42; Mismatches 187; Indels 31;
Qy 42 GTCVNCRPNFYINGAAQ-----GEANCNQPPAANNNARGI-----CVPQIQINHVGSVTIN 91	Qy 80 PCQINRVGSVTNAGDLATLATOQCTCOPGTGTAQDGVNDVFDRSAAOCVKCRPNFYINGGG 139
Db 1111 GVCIEAARGYFAVGAANTEQSVIAQGDTTGYTIANGNYKGIDACBESAPDATAAE 170	Db 898 QCKTKCPGTYIN-----CNPQNCKCOPKNTDKPRTDN-----EICKCNDDG 988
Qy 92 AGDLATLATQC-----STOCPTGT-----ALDD-----GV-----116	Db 140 SPQ-----GEAPGQVQFAAGAAAAGYANTTSQCVPCQLNNKNDSPATAGAQANLATOQC 191
Db 1711 AKVYAT-CRKCGSKYKLDRKVNVCDKAQCNSSGSTMKEVADDSSENGNMKCVSCSDNLNGVA 229	Db 952 TKQCTKDEAP-----CNPQNCKCOPKNTDKPRTDN-----EICKCNDDG 988
Qy 1177 ---TDVFDSEA---AQCVKCKPNFY-----NGGSPQE-----APGYQVFAAG 154	Qy 192 ----SNQ-CPTGTVLDDGTVLVENTSATL-----CVPKCRPNFYNGGSPQEAPGVQVFAAGAA 246
Db 230 NCDICSYDQSOKSKLKCCTCDNNYLKTTSEGTSVYKDDQSAGNMKCLPCN 289	Db 989 YLTPTPNQCVPDCTAISG-----YYGDDIKKCKACNPE-----CA 1022
Qy 155 AAAAGVAATISQCVPCQLNKNDSPA-----TAGAQANLAT----QC 191	Qy 247 AGYAAVTSQCVPCQINK-----NDSPATAGAQANLATOQCS 281
Db 290 DSTGCA-----RGATCALVSGREGAAAVLTCASACTDGYPKSADKTECEAVSNICKTPGCKAC 345	Db 1023 ECYGPANNQCTACPVGKMLQYTDTNTPVNGGTCMDQCSVSSTNDGCAEQIGGTAVCS 1082
Qy 192 SNOQPTGTWLD---DGVLIVVENTSATL-----215	Qy 282 ----TOCPT---GTATGTVLUVNSNSTQSOCTANYF-----FNGNFEAKRSQCL 326
Db 346 SNEGKENEVCTDCGSTYLTPTSOIDSACKIGNYYGATEGAKKLCKECATAANCKTCDDQ 405	Db 1083 KCXNTQOPLNGNCAASSRVAFCATISGACTKCNNEGYFLKDGGCYQTDRQPGHQVCSNA 1142
Qy 216 ---CVCKRPNFYINGG---SPOGEAPGVQVFAAGAAAAGVAVTSQCVPCQINK-----NDS 266	Qy 327 ----KCPSKSKTPAHAPGNTQATOQCLTQCPGTVLDGTSITNFVRAFATECTKCSAGF 381
Db 406 GQCQAQNDFGFKNGDAGCSQCHE-----SCKTCAGTA-----SDCTECPTGKALRYGDDG 455	Db 1143 QGGNGKQCTCANGLAASDGNCA -----ECHSICATCATDAD-----AADKCKTCTCAGY 1189
Qy 267 P-----ATAGAQANLATOQCS-----TOCPTGTAI-----ODGV-----TLVEVSNS 301	Qy 382 FASKTTGFTAGTDTCTCTKLTGSTGATAKYEAATORVQCASTT 425
Db 456 TKGTCGEGCCTGTGAGACTKGTLTDGASYCSEATTETYQONGVCAPKASRATPICNDS 515	Db 1190 YKENGDDTTAG-----LCKKCSCEKI-SGCK-----QCYSSS 1219
Qy 302 STQ---CSQCTANYF-FNGN-FEA-----GKSOCILKCPVSKTTPAHAPGNTQATOQCLT 352	RESULT 4
Qy 5116 P1QNGVCGTGCAADNYFMNGCGYETVKGPTVCSAPNGTCTOKAAGYKLSGT-----LTV 573	A4B44
Db 353 CPAGTVIDDGTSITNFVASECTRCOSAGE---ASKTGFETAGDTCCTECTKKLTSITSGATAK 410	variant-specific surface protein - Giardia lamblia (strain GS/M)
Db 574 CSEF-----CKECACTKCDASCTCNGAATICKACATG-----624	C;Species: Giardia lamblia
Qy 411 VYAEATORVQCAS 423	C;Accession: A4B44
Db 625 YKTTASGEAGACTS 637	C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 07-May-1999
RESULT 3	C;Accession: A4B434
T42017 cysteine rich protein - Giardia intestinalis	R;Nasir, T.E.; Mowatt, M.R.
C;Species: Giardia intestinalis	Mol. Bioch. Parasitol. 51, 219-228, 1992
C;Date: 03-Dec-1997 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000	A;Title: Characterization of a Giardia lamblia variant-specific surface protein (VSP
R;Chen, N.; Upcroft, P.; Upcroft, J.	A;Reference number: A4B434; MUID:92244292; PMID:1574080
Parasitology 111, 423-431, 1995	A;Accession: A4B434
	A;Status: preliminary
	A;Molecule type: mRNA
	A;Cross-references: GB:M80480; PID:9159142; PID:9159143
	Query Match 9.0%; Score 210.5; DB 2; Length 557;
	Best Local Similarity 22.9%; Pred. No. 2.7e-06;
	Matches 114; Conservative 48; Mismatches 184; Gaps 29;
Qy 7 LILIISFINELRAVECPDGTQTAQGLTD-----VGAADLGTCVNCRPNFYNGGAQGE 61	Qy 7 LILIISFINELRAVECPDGTQTAQGLTD-----VGAADLGTCVNCRPNFYNGGAQGE 61
Db 1 MELLINCLIASTLACG-----STQCNVCAEKCEVGETI-----CTQCKONVPINGVCEAA 55	Db 1 MELLINCLIASTLACG-----STQCNVCAEKCEVGETI-----CTQCKONVPINGVCEAA 55
Qy 62 ANGN-----QFPAANNAARGICVPCQ-----INRVGSVTNAQDLATL-----ATQCSQCPGPTGAL 112	Qy 62 ANGN-----QFPAANNAARGICVPCQ-----INRVGSVTNAQDLATL-----ATQCSQCPGPTGAL 112

Db	56 ASSNTKQKASADEASDQTGCKLSTTFMYKGGCYDKTGNLGRICKTEGSDAGKCGACN	115	Db	289 CSS---NNNSPLGIACTDCPAGTYA7SGDSCSVCKPCHNTCAAGCOTDDRETSCTACYP	344
Qy	113 DGVTDYFDRSAA----QCWKCKRNFYINGSPQGEAFQVOFAAGAA-AGVAAATSQ	166	Qy	271 -----GAQANLATAQCS-----TQCPGTTAQDGV-----	294
Db	116 DE--KGFDNDAANNVDSCTSC-----GDTGVTIPGSSTSKTKYGVAGAAKC	161	Db	345 GYSLLYESNGATGRVKETGAFITNCADGQCTANYGGAKYCTQCKDGYAPTDGTCATAV	404
Qy	167 CVPQLNKNDSPTA--GAQANL----ATOC--SNQCPGTIVLDGVTLYFNT	211	Qy	295 ----TLVFSNSSSHQCSQIANY----FFNQNFNEAGKSQQL-----RCPVRT-T	334
Db	162 TKPSQISENTGKTEATCNEANNTYLAKEYSSPMSATSCVSAAEICKTG-----YFT	212	Qy	295 ----TLVFSNSSSHQCSQIANY----FFNQNFNEAGKSQQL-----RCPVRT-T	334
Qy	212 SATL----CVKCRNFYINGSPQGEAFQVOFAAGAAAAGYAAVTSQCVPQINKND	266	Db	405 AAGRDVSYTATGSKCTACTGNYALSGGCYNTQTLPGKSYCKAVANSNDGKC--KTCA	461
Db	213 TDTTDSKKCLTC----STAKGGTDGSACELLSTTRASTVLSACSTN-NLS	264	Qy	335 PAHAGNTATQATQCLTICPAVTLDGDTSTNEVASATETKCSAGAFFASRTTGTAGTD	394
Qy	267 PATAGQAQNLTATCOSTCOPGTGTAQDGVTLYFENNSSPCQSCQIANYFFNGNEAGSQL	326	Db	462 NGQAPDPATNFCPICDSTCAECSTKND-----ADACTPKCFPYY--KIG----N	504
Db	265 P-----LNKEQMDCPAGTYADSNC--KPCHTSASCQKD-----NTE--SSCT	305	Qy	395 TCTBCTKKLTSQ	406
Qy	327 KCPVSKTPAHPAGNTATQATO----CLTTCPAGTVD--DGTSNTFVASATECTKCS	378	Db	505 KC-TRKTESSNNNG	516
Db	306 AC-----YPSVSLVSGTDNTKGTTAEC-TGKVLENCADGQCTATGSKCCK	355			
Qy	379 AGFF----ASKTGTFGTATGDFICTECKK--LTSGA-----407		RESULT 6		
Db	356 SGFVPPVNGLCVSAETARRAAPPGSTPDKTN-----VCTACTEKYFPLESGCYQAEKFPFG	409	A42125		
Qy	408 -TAKVYAETQKVOCAS	423	trophozoite cysteine-rich surface antigen 170 - Giardia lamblia		
Db	410 NTLCUTTADAGKCTTCAN	426	N; Alternative names: CRP170; cysteine-rich surface antigen CRP170		
			C; Species: Giardia lamblia		
			C; Date: 05-Dec-1998 #text_change 01-Dec-2000		
			C; Accession: A42125; S00530; S48056		
R; Mowatt, M.R.; Aggarwal, A.; Nash, T.E.	Mol. Biochem. Parasitol. 49, 215-227, 1991		R; Adam, R.D.; Yang, Y.M.; Nash, T.E.		
A; Reference number: A45664; PMID:92131058; PMID:1775165	A; Title: The cysteine-rich protein family of Giardia lamblia: loss of the CRP170		Mol. Cell. Biol. 12, 1194-1201, 1992		
A; Accession: A42125; MUID:92186830; PMID:1545800	A; Reference number: A42125; MUID:92186830; PMID:1545800		A; Reference: A42125; MUID:92186830; PMID:1545800		
A; Accession: A42125	A; Accession: A42125		A; Accession: A42125		
A; Residues: 1-98 <ADA1>	A; Residues: 1-98 <ADA1>		A; Residues: 1-98 <ADA2>		
A; Cross-references: GB:MB81937; NID:9159124	A; Cross-references: GB:MB81937; NID:9159124		A; Cross-references: GB:MB3933; NID:9159122		
A; Experimental source: trophozoite	A; Experimental source: trophozoite		A; Note: sequence extracted from NCBI backbone		
A; Note: the authors report but do not show 19 tandem repeats of the sequence of resid	A; Note: the authors report but do not show 19 tandem repeats of the sequence of resid		R; Adam, R.D.; Aggarwal, A.; Lal, A.A.; de la Cruz, V.F.; McCutchan, T.; Nash, T.B.		
A; Accession: B42125	A; Accession: B42125		J. Exp. Med. 167, 109-119, 1988		
A; Molecule type: DNA	A; Molecule type: DNA		A; Title: Antigenic variation of a cysteine-rich protein in Giardia lamblia		
A; Cross-references: EMBL:X06741; NID:919355; PID:929603	A; Cross-references: EMBL:X06741; NID:919355; PID:929603		A; Accession: S00530		
R; Yang, Y.; Adam, R.D.	R; Yang, Y.; Adam, R.D.		A; Molecule type: DNA		
Nucleic Acids Res. 22, 2102-2108, 1994	Nucleic Acids Res. 22, 2102-2108, 1994		A; Cross-references: EMBL:X06741; NID:919355; PID:929603		
A; Title: Allele-specific expression of a variant-specific surface protein (VSP) of G1	A; Title: Allele-specific expression of a variant-specific surface protein (VSP) of G1		R; Yang, Y.; Adam, R.D.		
A; Reference number: S48056; PMID:8029018	A; Reference number: S48056; PMID:8029018		Nucleic Acids Res. 22, 2102-2108, 1994		
A; Accession: S48056	A; Accession: S48056		A; Gene: VSP46		
C; Keywords: surface antigen; tandem repeat	C; Keywords: surface antigen; tandem repeat		C; Genetics:		
Query Match	Best Local Similarity	9.0%	Query Match	9.0%	Query Match
Matches 125; Conservative 50; Mismatches 188; Indels 189; Gaps 33;	Score 210.5; DB 2; Length 596;	Score 210.5; DB 2; Length 596;	Score 206; DB 2; Length 1766;	Score 206; DB 2; Length 1766;	Score 206; DB 2; Length 1766;
Qy	5 IILLIISLFINELAACPDCD-GPQTQALTD-VGAADI-----GTC-----44		Qy	76 GICVPOQINRVGSTVNAGDLATLATOQ----STQCPGT-GTADLGVTDFD-----RS 123	
Db	4 IAFYLILSTP----AVDKNSGNSEAGQCDTIGDTECMCQNGKVPIINGLCTAHSEE	58	Db	119 GYCGAC---KDGYKKNSDVAAT-ADSCIACEDANCATGGAGNKCTKICGYFYGATGN	174
Qy	45 ----VNCRPN-----FYYNGQ---GEANGN---QPFANNAAAR--75		Qy	124 AAQYKCK---PMFYIYNGG-----SPQGEAPGVQYFAAGAAAAGYAAVTSQCVP---C 170	
Db	59 AVTNAGCKKNGTNTESKVKCGCNGYFLHGGCYKIGEAGPNLICADEAASPQARTA	118	Db	175 EGGGCKCDTTGPNSYKGAGCARCEPKRNAPAKCIEADYLTEADEQTSVSEAVC	234
Qy	76 GICVPOQINRVGSTVNAGDLATLATOQ----STQCPGT-GTADLGVTDFD-----RS 123		Qy	171 QLNKNDSPTA--AGAQNLTATQCSNQCPGTGVLDDGV-----TLVFNSTAT-----LCYK	218
Db	119 GYCGAC---KDGYKKNSDVAAT-ADSCIACEDANCATGGAGNKCTKICGYFYGATGN	174	Db	235 REGTHFPTDSQGNGKVKCVSC-----GTTNNGGTENGECTSKESARAGTEITCTK	288
Qy	124 AAQYKCK---PMFYIYNGG-----SPQGEAPGVQYFAAGAAAAGYAAVTSQCVPQINNNDSPTA--270		Qy	219 CRPNFYIYNGGSPQGEA----PGVQVFAAGAAAAGVAAVTSQCVPQINNNDSPTA--	71
Db	1101 AVDC----QGSAYXTTDSAKECKKCNAPCTACAGTADKCKTCDANGAAAPZLKKTNP	1156	Db	1101 AVDC----QGSAYXTTDSAKECKKCNAPCTACAGTADKCKTCDANGAAAPZLKKTNP	1156

RESULT 9	A; Residues: 1-3084 <SARS>
	A; Cross-references: EMBL:J04064; NID:9309419; PIDN:AAA39410.1; PID:9309420
	A; Accession: A30449
	A; Molecule type: protein
	A; Residues: 183-195;70-571,'A',573-586;596-612,'X',614-617,'EMR';630-646;1217-1222,
	R; Hartl, L.; Oberbaumer, R.
	Eur. J. Biochem. 173, 629-635, 1988
	A; Title: The N terminus of laminin A chain is homologous to the B chains.
	A; Reference number: S00624; MUID:88225080; PMID:3267223
	A; Accession: S00624
	A; Molecule type: mRNA
	A; Residues: 1-208; 'T',210-334 <HAR>
	A; Cross-references: EMBL:X07737; NID:952857; PIDN:CAA30561.1; PID:952858
	A; Accession: A30450
	A; Molecule type: protein
	A; Residues: 311-335,'N',337-339;630-642,'D',644-692-734;737-748,'X',750-760,'G',762-763;
	3-1389;1443-1459 <HR2>
	A; Note: the sequence from Fig. 7 is inconsistent with that from Fig. 5 in Having 209
	R; Mann, K.; Deutzmann, R.; Timpl, R.
	Eur. J. Biochem. 178, 71-80, 1988
	A; Title: Characterization of proteolytic fragments of the laminin-nidogen complex and
	A; Reference number: S08895; MUID:89078415; PMID:2462498
	A; Accession: S08895
	A; Molecule type: mRNA
	A; Residues: 153-169 <MAN>
	R; Fujiwara, S.; Shinkai, H.; Deutzmann, R.; Paulsson, M.; Timpl, R.
	Biochem. J. 252, 453-461, 1988
	A; Title: Structure and distribution of N-linked oligosaccharide chains on various domains
	A; Reference number: S02678; MUID:88326259; PMID:2458101
	A; Accession: S02678
	A; Molecule type: protein
	A; Residues: 630-642,'D',644-692-734;737-748,'X',750-760,'G',762-763;
	R; Deutzmann, R.; Huber, J.; Schmitt, K.A.; Oberbaumer, I.; Hartl, L.
	Eur. J. Biochem. 177, 35-45, 1988
	A; Title: Structural study of long arm fragments of laminin. Evidence for repetitive C
	A; Reference number: S01190; MUID:89030693; PMID:3181157
	A; Accession: S01190
	A; Molecule type: mRNA
	A; Residues: 2538-3084 <DEU>
	A; Cross-references: EMBL:X13459; NID:955499; PID:9818014
	A; Accession: A30451
	A; Molecule type: protein
	A; Residues: 1911-1929;1997-2006;2033-2045,'X',2047-2054,'X',2056-2066,'X',2068-2105;2
	470;2487-2498;2502-2557;2581-2591,'X',2593-2594;2600-2610;2616-2645;2648-26
	93;2998-3005,'A',3007-3033,'V',3035;3068-3083 <DEE>
	A; Note: 2256-val was also found
	R; Olsen, D.; Nagayoshi, T.; Pazio, M.; Peltonen, J.; Jaakkola, S.; Sanborn, D.; Saksak
	Lab. Invest. 60, 772-782, 1989
	A; Title: Human laminin: cloning and sequence analysis of cDNAs encoding A, B1 and B2
	A; Reference number: A34961; MUID:89280632; PMID:2733383
	A; Accession: S14670
	A; Molecule type: protein
	A; Residues: 2424-2446;2440-2451;2461-2467;2487-2525;2557-2593;2600-2610;261
	-2942;2944-2964;2969-2976;2980-2993;2998-3000,'I',3002-3018,'V',3020-3034;3086-30
	C; Function:
	A; Description: interact with cells and with other basement membrane proteins to promote
	C; Keywords: laminin alpha-1 chain; laminin G repeat homology; laminin-type EGFR-like
	F;1-2a/Domain; signal sequence #status predicted <SG>
	F;25-27/Domain; VI <Dom6>
	F;27-33/Domain: laminin-type EGFR-like homology <LE01>
	F;33-40/1/Domain: laminin-type EGFR-like homology <LE02>
	F;40-45/8/Domain: laminin-type EGFR-like homology <LE03>
	F;46-50/7/Domain: laminin-type EGFR-like homology <LE04>
	F;51-51/9/Domain: laminin-type EGFR-like homology <LE04B>
	F;716-716/Domain: IIB <DO4B>
	F;716-746/Domain: laminin-type EGFR-like homology #status atypical <LE05>
	F;749-795/Domain: laminin-type EGFR-like homology <LE06>
RESULT 10	N; Alternative names: laminin chain precursor - mouse
	MIMMSA
	N; Alternative names: laminin chain A1
	C; Species: Mus musculus (house mouse)
	C; Date: 30-Jan-1991 #sequence-revision 30-Jun-1991 #text_change 19-Jan-2001
	C; Accession: A31771; A30449; S00624; A30404; S08895; S02678; S0190; A30451; S14670
	D; J. Biol. Chem. 263, 16536-16544, 1988
	A; Title: Laminin, a multidomain protein. The A chain has a unique globular domain and ho
	A; Reference number: A31771; MUID:89034134; PMID:3182802
	A; Molecule type: mRNA

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C;Accession: T25933
 R;Murray, J.; Wohldmann, P.
 R;Submitted to the EMBL Data Library, December 1996
 A;Description: The sequence of *C. elegans* cosmid W02C12.
 A;Accession number: Z20112
 A;Accession: T25933
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-1372 <MUR>
 A;Cross-references: EMBL:U80815; PIDN:AA837995.1; GSPDB:GN00022; CESP:W02C12.1
 A;Experimental source: strain Bristol N2; clone W02C12
 C;Genetics:
 A;Gene: CESP:W02C12.1
 A;Map position: 4
 A;Introns: 29/1; 66/1; 774/2; 823/2; 1046/1; 1108/2; 1298/1
 Query Match 8.3%; Score 195.5; DB 2; Length 1372;
 Best Local Similarity 23.9%; Pred. No. 5..3e-05;
 Matches 105; Conservative 43; Mismatches 155; Indels 137; Gaps 27;
 QY 30 QAGLTIDVAGDLTCVRNCRPNFYNGGAQGEA----NGRQPFAANNAARGICVPCQIN 84
 Db 923 QAG----QVVVRDL--CVPZAPGTHS--PATGECELCPTGEQPL---TATETECPFC--- 969
 QY 85 RVGSVTNAGDLATLATOCTCOPCTGTALDDGTYTDVFDRSAACQVKCKPNFYNGGSPOQE 144
 Db 970 APGQT-TASGAISGECDNCPCQHGDSDLSID-----CYTGSYY----OP--- 1013
 QY 145 APGVQVPAAGAAAAGVAATISOCPVPCOLNKNDSPATAGAQANLATOCNSQCPGTIVLDG 204
 Db 1014 -----SAGA-----FECIPCGICK---TTLESEFTSEDECEDECPCDGEQL--- 1050
 QY 205 VTLVFNNTSATCUCVKCRPNFYNGGSPOQEPGNQVFAAGAAAGVRAAVTSOCYPCQINKN 264
 Db 1051 -----SASGYCOPCOIGTYRSRGEE----NKKCYAC---- 1076
 QY 265 DSPATAGAQANLAT---QCST-QCPTGTAIQDGVTLFVNSNSTQCSQCJANTFFENGNEA 320
 Db 1077 --PPGTTERTMSTRBQQCNPKCKPGQFL-----YKETKNCQFPGRGTEQN--EE 1123
 QY 321 GKSQCLKCPVSKTTPAHAPGNTATQTCQLT---C-PAGTVY----LDDGTSTNFVYASA 371
 Db 1124 OBSTCKLCAPDHTTA-APG-ATAESQCFSTNQCATGEYNCSWHANCIDLDPDENDVS- 1178
 QY 372 TCTTKCSAGFFASKTGTGFTAGTDTC-----TEC-TKKLTSQATAYKAATQ 417
 Db 1179 YEC-RCPGYRNGTHTDACNDFCLNDGICKNNNIGVCEICKRDHFESSDRCELRFQASN 1237
 QY 418 KVOCASFFFAKFLSISLLFI 437
 Db 1238 NKLUWIAVTVGVVVVGIIIV 1257
 RESULT 12
 T42215
 zonadhesin - mouse
 N;Alternate names: sperm-specific membrane protein
 C;Species: Mus musculus (house mouse)
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999
 C;Accession: T42215
 R;Gao, Z.; Garbers, D.L.
 J. Biol. Chem. 273, 3415-3421, 1998
 A;Title: Species diversity in the structure of zonadhesin, a sperm-specific membrane
 A;Reference number: Z22080; PMID:98123114; PMID:9452463
 A;Accession: T42215
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-5376 <GAO>
 A;Cross-references: EMBL:U97068; PIDN: AAC26680.1
 C;Genetics:
 A;Gene: Zan
 A;Map position: 5
 RESULT 11
 T25933
 hypothetical protein W02C12.1 - *Caenorhabditis elegans*
 C;Species: *Caenorhabditis elegans*

C;Function:
A;Description: functions in multiple cell adhesion processes
C;Keywords: cell adhesion

Query Match	23 CPDGQTQAGLTDVGADLGTCV-----NCRPNFYNNGGAAQGEANGNQPF	Score 189; DB 2; Length 5376;	Qy 69 ANNA - ARGIC-----VPCQINFGSTVNAGDLATIATQC----- STQ 105
Best Local Similarity	21.2%; Pred. No. 0.00044;	Matches 56; Mismatches 204; Indels 152; Gaps 28;	Db 1268 CLNNAVCTNNICVSTFCSCSTNCVQISQCNQYVSVSQCQGSSQCLNSNCISSICQ 1327
Matches 111; Conservative 56;			Db 1328 CPGTQDQSNGCQPVQLINNQCYNTVSIQFQCFPOOGLNSQCNMSMC 1387
Qy 23 CPDGQTQAGLTDVGADLGTCV-----NCRPNFYNNGGAAQGEANGNQPF	69		Qy 106 CPTGTADDDGYTDVFDRSAACQ----- VRCK- 131
Db 2437 CPGTIVCKN-SNDGSNN-CVKTSLOCPAHSKFTDCUUPCHPSCSDPGHREGISTNA	2491		Db 1328 CPGTQDQSNGCQPVQLINNQCYNTVSIQFQCFPOOGLNSQCNMSMC 1387
Qy 70 ANNAARG-ICVP-----COINRYSV-----TNAG-----DLATLTAQ	101		Qy 132 -----PNEYFNGSPGPEAPGVQVFAAGAAAAGVAATVSOCPVCOLNKNDSPATAGA 183
Db 2492 HSNCRKCGCVCPGYVIRNDKCVLIECGQHTQGFIAGKMTSREGSOSCDCMEGYVIR	2551		Db 1388 QCPTSTTNVNGYCOGS-NGQCNQVNL-----INNOQY-----NTVSIGF 1427
Qy 102 CST-OCPGTG-ALDDGYTDVFDRSAAC-----VCKCPNFYNNGGSPQGBAPGVQF	151		Qy 184 QANLATOCSN-----QCPGTTVLDGVTLVFNITSATLVCKRNFNYNGGSPQGEA 234
Db 2552 CQNFCQPSGTYCQDFEDGTNSNCAN-ITLQCPAHSSFTNCLPPCQPSCSDEPHCGSSTK	2610		Db 1428 QCQFRQCLGNSQCLNSICQCPGSG-----SNVNGYCOGS-NGQCN 1467
Qy 152 AAGAAAAG-----VAATVSQCP-----CQ-----LNKNDSPATAGAQANL	187		Qy 235 PGVQWFAAGAAAAGVAATVSOCPVCPQLINKNDSPATAGAQANLATOC-----STQCP 285
Db 2611 APSACQEGCYCEPDIVVLLANKCVRGEGCKDAQGVLPADKTIWINKCTQTCACYVTG	2670		Db 1468 NSNOVY-----YNNQCY-----NTVTP-ITSQCCQITQQGLNSQCMNFCQCP 1508
Qy 188 ATQCSN-QCPTGTGTVLDDGVTLVFNITSATLVCV-----CRPNFYNNGGSPQGE	233		Qy 286 TGTAIQDGTVLVSNSNSSTQCS-----QCTANYFFNGMF----EAGKSQ-C 325
Db 2671 -HCRDFQCPGTYCKD-----TKDANSCTETILQCPDHSLYTHCLPSCLLSCSDPDSL	2723		Db 1509 SGTTNVNNNCFCTTSSSSNULCSAGTQVLDSSNQPICLVSTCPNNISGQRVCC 1568
Qy 234 APGQVFAANGAAAAGVAATVSOCPVCPQLINKNDSPATAGAQANLATOC-----	293		Qy 326 LKCPVSKTPAHAPGNTATQATQCLT-----CPAGTVLDDGT--STNEVASAATECKCSAG 380
Db 2724 CRGTSPEAPTSTCKRGCV-----CDPDVYLNSND-----KCVLRLBECGCKDAQGVLPAG	2771		Db 1569 RKCGTNSSPQVSKASQGVOVY-----CFTINSOCASGYCISNGACPNINNTICSTTGPCTPFTG 1624
Qy 294 VTLVFSNSSTOCSOCIANIFYFFNQF-----EAGKQCLCPVSKTPAHAPGNT	342		Qy 381 -----FFASAKTGTGFTAGTDTIC-----TECKKLTSGATAKYYAEATQVQCY 421
Db 2772 KTWI-NRGQTSCSMCGGAIQCONFKECPSEAYCQDMEDGNNSNCSTSPLQ-CPHA-----	2823		Db 1625 QISVGQGCFNSYNG-----DRCRSQECLGSGQCNUNLQCPNS-FANVNCQCA 1674
Qy 343 ATQATOCLTTC-PAGTVL-----DGTSTNFVASAATECKCSAGFF-----ASK	385		RESULT 14
Db 2824 -SHYTNCPLTCOPSCSDPGEFESSKAPBACKECCVCPDFYVMLNNCKVPRTECGIRD	2882		T43251
Qy 386 TTGFTAGTDTCCTKLTTSGATAKYYAEATQVQCASTTEFAK	428		furin (SC 3.4.21.75) - fall armyworm
Db 2883 TQGYLIPPAK-TWINRGCTQSCCTCRGAIQCQKYHCSSGTYCK 2924			N;Alternate names: Paired basic amino acid cleaving enzyme; proteinase; convertase; se
RESULT 13			C;Species: Spodoptera frugiperda (fall armyworm)
T15881 hypothetical protein D1044.3 - Caenorhabditis elegans			C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
C;Species: Caenorhabditis elegans			R;Accession: T43251
C;Accession: T15881			C;Accession: EMBL:268888; NID:g1167859; PID:e219690; PIDN:CAA93116.1
R;Pauley, A.			A;Cross-references: EMBL:268888; NID:g1167859; PID:e219690; PIDN:CAA93116.1
submitted to the EMBL Data Library, June 1994			A;Experimental source: clone Sfurin 6; ovary
A;Description: The sequence of <i>C. elegans</i> cosmid D1044.			C;Function:
A;Accession: T15881			A;Description: responsible for the endoproteolytic processing of proteins with spe
A;Status: preliminary; translated from GB/EMBL/DBJ			C;Keywords: hydrolase; serine proteinase
A;Molecule type: DNA			
A;Residues: 1-1895 <PAU>			Query Match
A;Cross-references: EMBL:U00065; NID:g495681; PID:9495684; PIDN:AAA50735.1; CESP:D1044.3			Best Local Similarity 7.8%; Score 183; DB 2; Length 1299;
C;Genetics:			A;Status: preliminary; translated from GB/EMBL/DBJ
A;Gene: CESP:D1044.3			A;Molecule type: mRNA
A;Introns: 35/2, 63/2; 150/1; 191/3; 207/2; 405/3; 551/3; 629/1; 764/1; 933/1; 959/2; 10			A;Residues: 1-1299 <CIE>
Query Match	7.8%; Score 183.5; DB 2; Length 1895;		A;Cross-references: EMBL:T268888; NID:g1167859; PID:e219690; PIDN:CAA93116.1
Best Local Similarity 20.7%; Pred. No. 0.00044;			A;Experimental source: strain Bristol N2
Matches 111; Conservative 46; Mismatches 171; Indels 209; Gaps 29;			A;Accession: T15881
Qy 23 CPDGQTQAG-LTDVGADLGTCV-----NCRPN-----FVYNGAAGEANGNQPF	68		A;Description: Cloning and functional characterization of FURIN from <i>Spodoptera frugi</i>
Db 1209 CVGGAMCNSTGTCQCTNGATAMYGYCISSSSSNSNOVINGMCYNTVQGGSCSFQSQ-	1267		A;Accession number: z22368
Qy 192 -----SNQCPGTTGTVLDDGVTLVFNITSATLVCKRNFNYNGGSPQGEAPGVQY	239		A;Accession: T43251
Db 804 -----CATCSERADGTSCEHHLVLDGTCMASCPSH-----YEEDDDCAKC- 847			A;Status: preliminary; translated from GB/EMBL/DBJ
Qy 132 PNFYNNGGSPQGAPGVQYFANGAAAAGVAATVSOCPVCOLNKNDSPATAGAQANLATQ	191		A;Molecule type: mRNA
Db 748 PCHESECTCVGPGDSCLTCAPIHLLADLAYCLQQ-PDGWE---DSEASVCRCAAH 803			A;Residues: 1-1299
Qy 72 NAARGICVCPQINFVSVNAGDLATIATQCSTQCPGTTDGRVTDVDRSAACQVCK 131			A;Cross-references: EMBL:T268888; NID:g1167859; PID:e219690; PIDN:CAA93116.1
Db 804 -----CATCSERADGTSCEHHLVLDGTCMASCPSH-----YEEDDDCAKC- 847			A;Experimental source: clone Sfurin 6; ovary
Qy 848 -----HESECTCVGPGDSCLTCAPIHLLADLAYCLQQ-PDGWE---ETQCVTCH-----PSTALDGRCVTSC 880			A;Accession: R;Accession: T43251
Qy 192 -----SNQCPGTTGTVLDDGVTLVFNITSATLVCKRNFNYNGGSPQGEAPGVQY	239		A;Description: responsible for the endoproteolytic processing of proteins with spe

Search completed: February 11, 2003, 19:48:39
 Job time : 28.0286 secs

RESULT 15

A48579 trophozoite surface protein TSP11 - Giardia lamblia

C;Species: Giardia lamblia
 C;Date: 01-Dec-1993 #sequence-revision 18-Nov-1994 #text_change 29-Jan-1999
 C;Accession: A48579
 R;EY, P.L.; Khamma, K.K.; Manning, P.A.; Mayrhofer, G.
 Mol. Biochem. Parasitol. 58, 247-257, 1993
 A;Title: A gene encoding a 69-kilodalton major surface protein of Giardia intestinalis
 A;Reference number: A48579; MUID:9324125; PMID:8479449
 A;Contents: Ad-1
 A;Accession: A48579
 A;Status: preliminary
 A;Molecule type: nucleic acid
 A;Residues: 1-667 <EY1>
 A;Cross-references: GB:M95814; NID:gi159106; PID:gi159107
 A;Note: sequence extracted from NCBI backbone (NCBIPN:130056, NCBIPN:130058)

Query	7 LILLISLFINELRAVPCPVGTTQ-----QAGLTDV-----GAADL-GHCVNCRPNFV	52
Db	1 MILLAIFYFYFLISTLTAKCTTQTQTCCEAKCCEMVEGETICTRCOTGKPIDGKCVDATAN-	58
Query	53 YNGGAQGEANGNQPFANNAARG--ICVPCQINRGSVT-----NAGDLATLATQC	102
Db	59 ANCKNSGDDANQVCKNSVSPNTLCTVSPDGVCSSVAANEYFVPPNADATHDSVVSC	118
Query	103 STQCPGTAQDDGVDFRSAROVCVKCPNFYNGGSPQGPQYQEAAAGVAA	162
Db	119 SEETPIHLANKXIGV----AGCATC-----SAP----RAGDENTPRAT-----	157
Query	163 VTSQCPVCPQLNKNDSPATAGAQANLATQCSNQCPGTGLDDGTVLFNTSAT-----LCV	217
Db	158 ----CPKCAAGFLHPSEG-----LSSEETCPEG-----YFHTAYAESKKTK	198
Query	218 KCRPNFYYNGGSPQGPQGPQYQVFAAAGAAAGVAAVTSSQVCPQINKNDSPATAGAQANLA	277
Db	199 SC-----TGGS---SEAPNPK-----GI----GDOLKCMYNE-----ASGNTL	229
Query	278 T-----QCSTQ-----OPTGAIQDQVTLVFSNSSTQ-----	304
Db	230 TCEBKCSAQKPSLDKTSNDCTGQNCAFCSSGGDCEGCDSG-FILDSQNCVKSDDCTEN	288
Query	305 -----CSOCIANYFFN-----GNPEAG-----KSQCLKCPVS-K	332
Db	289 CKACTNPRAAEVCTE-TISTHLTPISQCYQCOALGNYFAGTNADNVKACKETVANCK	348
Query	333 T-----TPAHAGNT-----ATQTOCLTCAGTVL-----DDGTSTN	366
Db	349 TCNDQGOCQTCDGFYKNGDACPSCHECSCTCSAGTASDC-TECPTGKALKYGNDDGTK-	405
Query	367 FVASATECTKCSAGFFASKTGTFTA GTDCTCTEKKLTSGATAKVYAPATOYQCASTT	425
Db	406 --GTCGESETTGUGSGACKTCSGTTGASYCSECDTQ-----NEYPQNGICTSTT	453